

[illegible]

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Gene Shears Pty. Limited
(B) STREET: Suite 1, Building 5, 105 Delhi Road
(C) CITY: North Ryde
(D) STATE: North Ryde
(E) COUNTRY: Australia
(F) POSTAL CODE (ZIP): NSW 2113

(A) NAME: PAUL, Wyatt
(B) STREET: c/o Nickerson Biocem Ltd, Cambridge Science
Park, Milton Rd
(C) CITY: Cambridge
(D) STATE: Cambridge
(E) COUNTRY: UK
(F) POSTAL CODE (ZIP): CB4 5GZ

(A) NAME: PEREZ, Pascaul
(B) STREET: c/o Biogemma, Campus Universitaire des
Cezeaux
(C) CITY: 24 Avenue des Landais
(D) STATE: Aubiere
(E) COUNTRY: France
(F) POSTAL CODE (ZIP): 63170

(A) NAME: HUTTNER, Eric
(B) STREET: c/o Groupe Limagrain Pacific Pty Ltd, GPO Box
475
(C) CITY: Canberra
(D) STATE: Canberra, ACT
(E) COUNTRY: Australia
(F) POSTAL CODE (ZIP): 2601

(A) NAME: BETZNER, Andreas Stefan
(B) STREET: Groupe Limagrain Pacific Pty Ltd, GPO Box 475
(C) CITY: Canberra
(D) STATE: Canberra, ACT
(E) COUNTRY: Australia
(F) POSTAL CODE (ZIP): 2601

(ii) TITLE OF INVENTION: Protein Complementation

(iii) NUMBER OF SEQUENCES: 68

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB98/00542

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 9..344
- (D) OTHER INFORMATION: /product= "Barnase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCTAGACC ATG GCA CAG GTT ATC AAC ACG TTT GAC GGG GTT GCG GAT TAT	50
Met Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr	
1 5 10	
CTT CAG ACA TAT CAT AAG CTA CCT GAT AAT TAC ATT ACA AAA TCA GAA	98
Leu Gln Thr Tyr His Lys Leu Pro Asp Asn Tyr Ile Thr Lys Ser Glu	
15 20 25 30	
GCA CAA GCC CTC GGC TGG GTG GCA TCA AAA GGG AAC CTT GCA GAC GTC	146
Ala Gln Ala Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val	
35 40 45	
GCT CCG GGG AAA AGC ATC GGC GGA GAC ATC TTC TCA AAC AGG GAA GGC	194
Ala Pro Gly Lys Ser Ile Gly Gly Asp Ile Phe Ser Asn Arg Glu Gly	
50 55 60	
AAA CTC CCG GGC AAA AGC GGA CGA ACA TGG CGT GAA GCG GAT ATT AAC	242
Lys Leu Pro Gly Lys Ser Gly Arg Thr Trp Arg Glu Ala Asp Ile Asn	
65 70 75	
TAT ACA TCA GGC TTC AGA AAT TCA GAC CGG ATT CTT TAC TCA AGC GAC	290
Tyr Thr Ser Gly Phe Arg Asn Ser Asp Arg Ile Leu Tyr Ser Ser Asp	
80 85 90	
TGG CTG ATT TAC AAA ACA ACG GAC CAT TAT CAG ACC TTT ACA AAA ATC	338
Trp Leu Ile Tyr Lys Thr Thr Asp His Tyr Gln Thr Phe Thr Lys Ile	
95 100 105 110	
AGA TAA	344
Arg	

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 112 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln
1 5 10 15
Thr Tyr His Lys Leu Pro Asp Asn Tyr Ile Thr Lys Ser Glu Ala Gln
20 25 30
Ala Leu Gly Trp Val Ala Ser Lys Gly Asn Leu Ala Asp Val Ala Pro
35 40 45
Gly Lys Ser Ile Gly Gly Asp Ile Phe Ser Asn Arg Glu Gly Lys Leu
50 55 60
Pro Gly Lys Ser Gly Arg Thr Trp Arg Glu Ala Asp Ile Asn Tyr Thr
65 70 75 80
Ser Gly Phe Arg Asn Ser Asp Arg Ile Leu Tyr Ser Ser Asp Trp Leu
85 90 95
Ile Tyr Lys Thr Thr Asp His Tyr Gln Thr Phe Thr Lys Ile Arg
100 105 110

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..18
- (D) OTHER INFORMATION: /note= "Figure 1A: B1 primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CATGGTCTAG AGTACTTG

18

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..16
- (D) OTHER INFORMATION: /note= "Figure 1A: B4 primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CCAGCCGAGG GCTTGT

16

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..16
- (D) OTHER INFORMATION:/note= "Figure 1A: B2 primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GCATCAAAAG GGAACC

16

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 228 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..228
- (D) OTHER INFORMATION:/note= "Figure 1B: Intergenic Sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGAAAAAAC GGCTTCCTGC GGAGGCCGTT TTTTTCAGCT TTACATAAAG TGTGTAATAA	60
ATTTTCTTTC AAACCTCTGAT CGGTCAATTT CACTTTCGGG ATCCGGTCCA ATCTGCAGCC	120
GTCCGAGACA GGAGGACATC GTCCAGCTGA AACCGGGGCA GAATCCGGCC ATTTCTGAAG	180
AGAAAAATGG TAACTGATA GAATAAAATC ATAAGAAAGG AGCCGCAC	228

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 323 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:1..273

(D) OTHER INFORMATION:/product= "Barstar"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATG AAA AAA GCA GTC ATT AAC GGG GAA CAA ATC AGA AGT ATC AGC GAC 48
Met Lys Lys Ala Val Ile Asn Gly Glu Gln Ile Arg Ser Ile Ser Asp
115 120 125

CTC CAC CAG ACA TTG AAA AAG GAG CTT GCC CTT CCG GAA TAC TAC GGT 96
Leu His Gln Thr Leu Lys Lys Glu Leu Ala Leu Pro Glu Tyr Tyr Gly
130 135 140

GAA AAC CTG GAC GCT TTA TGG GAT TGT CTG ACC GGA TGG GTG GAG TAC 144
Glu Asn Leu Asp Ala Leu Trp Asp Cys Leu Thr Gly Trp Val Glu Tyr
145 150 155 160

CCG CTC GTT TTG GAA TGG AGG CAG TTT GAA CAA AGC AAG CAG CTG ACT 192
Pro Leu Val Leu Glu Trp Arg Gln Phe Glu Gln Ser Lys Gln Leu Thr
165 170 175

GAA AAT GGC GCC GAG AGT GTG CTT CAG GTT TTC CGT GAA GCG AAA GCG 240
Glu Asn Gly Ala Glu Ser Val Leu Gln Val Phe Arg Glu Ala Lys Ala
180 185 190

GAA GGC TGC GAC ATC ACC ATC ATA CTT TCT TAA TACGATCAAT GGGAGATGAA 293
Glu Gly Cys Asp Ile Thr Ile Ile Leu Ser
195 200

CAATATAGAT CCCCCGGGCT GCAGGAATTC 323

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 91 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Lys Lys Ala Val Ile Asn Gly Glu Gln Ile Arg Ser Ile Ser Asp
1 5 10 15

Leu His Gln Thr Leu Lys Lys Glu Leu Ala Leu Pro Glu Tyr Tyr Gly
20 25 30

Glu Asn Leu Asp Ala Leu Trp Asp Cys Leu Thr Gly Trp Val Glu Tyr
35 40 45

Pro Leu Val Leu Glu Trp Arg Gln Phe Glu Gln Ser Lys Gln Leu Thr
50 55 60

Glu Asn Gly Ala Glu Ser Val Leu Gln Val Phe Arg Glu Ala Lys Ala
 65 70 75 80

Glu Gly Cys Asp Ile Thr Ile Ile Leu Ser
 85 90

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "Figure 1C: B3 primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TAATACGATC AATGGGAGAT G

21

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 194 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:9..194
- (D) OTHER INFORMATION:/note= "Figure 1D"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TCTAGACC ATG GCA CAG GTT ATC AAC ACG TTT GAC GGG GTT GCG GAT TAT
 Met Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr
 95 100 105

50

CTT CAG ACA TAT CAT AAG CTA CCT GAT AAT TAC ATT ACA AAA TCA GAA
 Leu Gln Thr Tyr His Lys Leu Pro Asp Asn Tyr Ile Thr Lys Ser Glu
 110 115 120

98

GCA CAA GCC CTC GGC TGG ATG GGC GGT GGC GGT TCC GGT GGC GGT GGC
 Ala Gln Ala Leu Gly Trp Met Gly Gly Gly Gly Ser Gly Gly Gly Gly
 125 130 135

146

AGC GGC GGC GGT GGT AGC GGG ATC CCC GGG TAC GGT CAG TCC CTT ATG
 Ser Gly Gly Gly Gly Ser Gly Ile Pro Gly Tyr Gly Gln Ser Leu Met
 140 145 150

194

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 62 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Ala Gln Val Ile Asn Thr Phe Asp Gly Val Ala Asp Tyr Leu Gln
 1 5 10 15

Thr Tyr His Lys Leu Pro Asp Asn Tyr Ile Thr Lys Ser Glu Ala Gln
 20 25 30

Ala Leu Gly Trp Met Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
 35 40 45

Gly Gly Gly Ser Gly Ile Pro Gly Tyr Gly Gln Ser Leu Met
 50 55 60

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 526 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..526
- (D) OTHER INFORMATION: /note= "Figure 1E"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TCTAGACCAT GCAGATCTTC GTGAAAACCT TGACCGGCAA GACCATCACT CTCGAGGTCG 60
 AGAGCAGCGA CACCATCGAC AATGTCAAGG CCAAGATCCA AGACAAAGAA GGTATCATTC 120
 TTCCTCACTC AATCTGGATT CTTCTCTTTA GCTTTTGTAA ATTGAGATCT CTTATCATTT 180
 ACTTGTTTCT CCTTTAAGGA ATCCCTCCGG ATCAGCAGAG ATTGATCTTC GCCGGAAGC 240
 AGCTCGAAGA TGGCCGTACT TTGGCTGACT ACAACATCCA GAAAGGTACG AAATCATCCG 300
 AATCCTTCTG TTGATCATTT CGATGATCTG ATTGTATAAA CTCTAATGGA TTGTTATCAT 360
 TTGTAAACAG AATCTACACT TCATCTTGTG TTGAGGCTTA GAGGTGGAGC ACAGGTTATC 420
 AACACGTTTG ACGGGGTTGC GGATTATCTT CAGACATATC ATAAGCTACC TGATAATTAC 480

ATTACAAAAT CAGAAGCACA AGCCCTCGGC TGGATGTAGA GGATCC

526

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 631 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION:1..631
 (D) OTHER INFORMATION:/note= "Figure 1F"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCTAGACCAT GCAGATCTTC GTGAAAACCT TGACCGGCAA GACCATCACT CTCGAGGTCG 60
AGAGCAGCGA CCATCGACAA TGTCAAGGCC AAGATCCAAG ACAAAGAAGG TATCATTCTT 120
CCTCACTCAA TCTGGATTCT TCTCTTTAGC TTTTGAAAT TCAGATCTCT TATCATTTAC 180
TTGTTTCTCC TTTAAGGAAT CCCTCCGGAT CAGCAGAGAT TGATCTTCGC CGGAAAGCAG 240
CTCGAAGATG GCCGTACTTT GGCTGACTAC AACATCCAGA AAGGTACGAA ATCATCCGAA 300
TCCTTCTGTT GATCATTTTCG ATGATCTGAT TGTATAAACT CTAATGGATT GTTATCATTT 360
GTAAACAGAA TCTACACTTC ATCTTGTGTT GAGGCTTAGA GGTGGAGCAT CAAAAGGGAA 420
CCTTGCAGAC GTCGCTCCGG GGAAAAGCAT CGGCGGAGAC ATCTTCTCAA ACAGGGAAGG 480
CAAACCTCCG GGCAAAGCG GACGAACATG GCGTGAAGCG GATATTAAT ATACATCAGG 540
CTTCAGAAAT TCAGACCGGA TTCTTTACTC AAGCGACTGG CTGATTTACA AAACAACGGA 600
CCATTATCAG ACCTTTACAA AAATCAGATA A 631

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION:1..20
 (D) OTHER INFORMATION:/note= "Figure 1G: B5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CACAAGTACT CTAGACCATG

20

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..19
- (D) OTHER INFORMATION:/note= "Figure 1G: B6"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CATCCAGCCG AGGGCTTGT

19

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..16
- (D) OTHER INFORMATION:/note= "Figure 1G: B7"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGCGGTGGCG GTTCCG

16

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..23
- (D) OTHER INFORMATION:/note= "Figure 1G: B8"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCACTAGTTC TAGAGTACTT GTG

23

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..18
- (D) OTHER INFORMATION:/note= "Figure 1G: B9"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCACAGGTTA TCAACACG

18

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..31
- (D) OTHER INFORMATION:/note= "Figure 1G: B10"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

GCGGATCCTC TACATCCAGC CGAGGGCTTG T

31

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature

(B) LOCATION:1..16
(D) OTHER INFORMATION:/note= "Figure 1G: B11"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GCATCAAAAG GGAACC

16

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:1..17
(D) OTHER INFORMATION:/note= "Figure 1G: B12"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGTCTAGAGT ACTTG TG

17

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:1..30
(D) OTHER INFORMATION:/note= "Figure 1G: Ubql6F"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GCTCTAGACC ATGCAGATCT TCGTGAAAAC

30

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..25
- (D) OTHER INFORMATION:/note= "Figure 1G: UbqlR"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CTGGATCCAC CTCTAAGCCT CAACA

25

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..24
- (D) OTHER INFORMATION:/note= "Figure 1G: Ubqla"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

TATGGATCCC CCGGGCTGCA GGAA

24

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "Figure 1G: Ubqlb"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

TCCACCTCTA AGCCTCAACA C

21

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..23
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 1, Primer RNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GCGGATCCAT GAAGGAGACC GCC

23

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..56
- (D) OTHER INFORMATION:/note= "Fig 3A; lane 2, RNI"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GCGGATCCAT GAAGGAGACC GCCGCCGCCA AGTTCGAGCG CCAGCACATG GACAGC

56

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..22
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 3, RN1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CATAGATCTT TAGCTGTCCA TG

22

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:1..28

(D) OTHER INFORMATION:/note= "Fig 3A: lane 4, Primer
RN-d"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CCAGATCTAT GAGCTCCTCC AACTACTG

28

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:1..63

(D) OTHER INFORMATION:/note= "Fig 3A: lanes 2/5, RNII"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TAAAGATCTA TGAGCACCTC CGCCGCCAGC TCCTCCAACCT ACTGCAACCA GATGATGAAG

60

TCT

63

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:1..21

(D) OTHER INFORMATION:/note= "Fig 3A: lane 6, RN2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

TCAGGTTTCCT AGACTTCATC A

21

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 59 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION:1..59
 (D) OTHER INFORMATION:/note= "Fig 3A, lanes 5/7, RNIII"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

AGGAACCTGA CCAAGGACAG GTGCAAGCCA GTCAACACCT TCGTCCACGA GAGCCTGGC

59

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION:1..19
 (D) OTHER INFORMATION:/note= "Fig 3A: lane 8, RN3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CTGGACATCG GCCAGGCTC

19

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION:1..48
 (D) OTHER INFORMATION:/note= "Fig 3A, lanes 7/9, RN IV"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CGATGTCCAG GCCGTCTGCA GCCAGAAGAA CGTGGCCTGC AAGAACGG

48

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 10, RN 4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

AGTTGGTCTG ACCGTTCTTG C

21

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..60
- (D) OTHER INFORMATION:/note= "Fig 3A: lanes 9/11, RN V"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

TCAGACCAAC TGCTACCACT CCTACAGCAC CATGTCCATC ACCGACTGCC GCGAGACCGG

60

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..19
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 12, RN5"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CTTGCTGGAG CCGGTCTCG

19

(2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 55 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..55
- (D) OTHER INFORMATION:/note= "Fig 3A: lanes 11/13, RN VI"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

CTCCAGCAAG TACCCTAACT GCGCCTACAA GACCACCCAG GCCAACAAGC ACATC

55

(2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 14, RN 6"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

CAGGCAACAA TGATGTGCTT G

21

(2) INFORMATION FOR SEQ ID NO: 40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..24
- (D) OTHER INFORMATION:/note= "Fig 3A: lane 15, Primer RN-b"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

CGGGATCCTT TAGACGGAGG CGTC

24

(2) INFORMATION FOR SEQ ID NO: 41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 66 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..66
- (D) OTHER INFORMATION:/note= "Fig 3A: lanes 13/16, RN VII"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

ATTGTTGCCT GCGAGGGTAA CCCTTACGTG CCTGTCCACT TCGACGCCTC CGTCTAAAGG

60

ATCCCG

66

(2) INFORMATION FOR SEQ ID NO: 42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..23
- (D) OTHER INFORMATION:/note= "Fig 3B: lane 1, PCR Primer RN_a"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

GCGGATCCAT GAAGGAGACC GCC

23

(2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 56 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..56
- (D) OTHER INFORMATION:/note= "Fig 3B, lane 2, RN I"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GCGGATCCAT GAAGGAGACC GCCGCCGCCA AGTTCGAGCG CCAGCACATG GACAGC

56

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..18
- (D) OTHER INFORMATION:/note= "Fig 3B, lane 3, RN 7"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

CCACCGCCGC TGTCCATG

18

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..57
- (D) OTHER INFORMATION:/note= "Fig 3B: lanes 2/4, RN VIII"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

GGCGGTGGCG GTTCCGGTGG CGGTGGCAGC GCGGGCGGTG GTAGCAAGAT CTTCGGG

57

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..18
- (D) OTHER INFORMATION:/note= "Fig 3B: lane 5, RN c"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

CCCGAAGATC TTGCTACC

18

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..45
- (D) OTHER INFORMATION:/note= "Fig 4, lane 1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

AAAGAGACAG CAGCCGCAAA GTTTGAGCGT CAGCATATGG ATAGT

45

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Peptide
- (B) LOCATION:1..32
- (D) OTHER INFORMATION:/note= "Fig 4A: lane 2"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION:17
- (D) OTHER INFORMATION:/note= "'Xaa" = corresponds to an ochre stop codon (UAA)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Xaa Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION:1..63
(D) OTHER INFORMATION:/note= "Fig 4A: lane 3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

GGATCCATGA AGGAGACCGC CGCCGCCAAG TTCGAGCGCC AGCACATGGA CAGCTAAAGA 60
TCT 63

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 106 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
(B) LOCATION:1..106
(D) OTHER INFORMATION:/note= "Fig 4A: lane 4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

GGATCCATGA AGGAGACCGC CGCCGCCAAG TTCGAGCGCC AGCACATGGA CAGCGGCGGT 60
GGCGGTTCCG GTGGCGGTGG CAGCGGCGGC GGTGGTAGCA AGATCT 106

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:1..330

(D) OTHER INFORMATION:/note= "Fig 4B: lane 1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

AGCACCAGTG CTGCCAGTTC TTCCAACACTAC TGTAACCAGA TGATGAAGTC TAGAAACTTG	60
ACCAAGGACA GATGTAAGCC AGTTAACACA TTTGTCCACG AGAGTTTGGC TGATGTCCAA	120
GCCGTCTGCA GTCAGAAAAA CGTTGCATGC AAGAACGGTC AAACGAACTG TTACCAGAGT	180
TACAGCACCA TGTCCATCAC TGA CTGTCGT GAGACAGGCT CGAGCAAGTA TCCTAATTGT	240
GCTTACAAGA CCACACAGGC GAACAAACAC ATCATTGTTG CTTGTGAAGG TAACCCTTAC	300
GTTCTGTCC ACTTTGACGC CAGTGTTTAA	330

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 132 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION:1..132

(D) OTHER INFORMATION:/note= "Fig 4B: lane 2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

Met	Ser	Thr	Ser	Ala	Ala	Ser	Ser	Ser	Asn	Tyr	Cys	Asn	Gln	Met	Met
1				5					10					15	
Lys	Ser	Arg	Asn	Leu	Thr	Lys	Asp	Arg	Cys	Lys	Pro	Val	Asn	Thr	Phe
			20					25					30		
Val	His	Glu	Ser	Leu	Ala	Asp	Val	Gln	Ala	Val	Cys	Ser	Gln	Lys	Asn
		35					40					45			
Val	Ala	Cys	Lys	Asn	Gly	Gln	Thr	Asn	Cys	Tyr	Gln	Ser	Tyr	Ser	Thr
	50					55					60				
Met	Ser	Ile	Thr	Asp	Cys	Arg	Glu	Thr	Gly	Ser	Ser	Lys	Tyr	Pro	Asn
65					70				75					80	
Cys	Ala	Tyr	Lys	Thr	Thr	Gln	Ala	Asn	Thr	Asp	Cys	Arg	Glu	Thr	Gly
				85					90					95	

Ser Ser Lys Tyr Pro Asn Cys Ala Tyr Lys Thr Thr Gln Ala Asn Lys
100 105 110
His Ile Ile Val Ala Cys Glu Gly Asn Pro Tyr Val Pro Val His Phe
115 120 125
Asp Ala Ser Val
130

(2) INFORMATION FOR SEQ ID NO: 53:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 346 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION:1..330
(D) OTHER INFORMATION:/note= "Fig 4B, lane 3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

AGATCTATGA GCACCTCCGC CGCCAGCTCC TCCAACACT GCAACCAGAT GATGAAGTCT	60
AGGAACCTGA CCAAGGACAG GTGCAAGCCA GTCAACACCT TCGTCCACGA GAGCCTGGCC	120
GATGTCCAGG CCGTCTGCAG CCAGAAGAAC GTGGCCTGCA AGAACGGTCA GACCAACTGC	180
TACCAGTCCT ACAGCACCAT GTCCATCACC GACTGCCGCG AGACCGGCTC CAGCAAGTAC	240
CCTAACTGCG CCTACAAGAC CACCCAGGCC AACAAGCACA TCATTGTTGC CTGCGAGGGT	300
AACCCCTTACG TGCCTGTCCA CTTGACGCC TCCGTCTAAA GGATCC	346

(2) INFORMATION FOR SEQ ID NO: 54:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 331 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION:1..331
(D) OTHER INFORMATION:/note= "Fig 4B, lane 4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

AGATCTATGA GCTCCTCCAA CTA CTGCAAC CAGATGATGA AGTCTAGGAA CCTGACCAAG	60
--	----

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 270 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..270
- (D) OTHER INFORMATION:/note= "Fig 4c iii"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ATGAAGAATG TTTTAGTAAG GTCAGCTGCG CGAGCTCTGC TTGGCGGCGG TGGGCGGAGC	60
TACTACCGCC AGCTCTCAAC GCGGCGGATC GTGGAACAGA GACACCAGCA CGGTGGCGGC	120
GCGTTTGGA GCTTCCACTT AAGAAGGATG AAGGAGACCG CCGCCGCCAA GTTCGAGCGC	180
CAGCACATGG ACAGCGGCGG TGGCGGTTCC GGTGGCGGTG GCAGCGGCGG CGGTGGTAGC	240
GGGATCCCCG GGTACGGTCA GTCCCTTATG	270

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 465 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..465
- (D) OTHER INFORMATION:/note= "Fig 4C iv"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

ATGAAGAATG TTTTAGTAAG GTCAGCTGCG CGAGCTCTGC TTGGCGGCGG TGGGCGGAGC	60
TACTACCGCC AGCTCTCAAC GCGGCGGATC GTGGAACAGA GACACCAGCA CGGTGGCGGC	120
GCGTTTGGA GCTTCCACTT AAGAAGGATG AGCTCCTCCA ACTACTGCAA CCAGATGATG	180
AAGTCTAGGA ACCTGACCAA GGACAGGTGC AAGCCAGTCA ACACCTCCGT CCACGAGAGC	240
CTGGCCGATG TCCAGGCCGT CTGCAGCCAG AAGAACGTGG CCTGCAAGAA CGGTCAGACC	300
AACTGCTACC AGTCCTACAG CACCATGTCC ATCACCGACT GCCGCGAGAC CGGCTCCAGC	360

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

GGTGGATCCA GCTCCAATA CTGCAAC

27

(2) INFORMATION FOR SEQ ID NO: 61:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

CGGGATCCTT AGACGGAGGC GTCG

24

(2) INFORMATION FOR SEQ ID NO: 62:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

GTCCTTAAGA AGGATGAGCT CCTCCAATA C

31

(2) INFORMATION FOR SEQ ID NO: 63:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

CGGGATCCTT AGACGGAGGC GTCG

24

(2) INFORMATION FOR SEQ ID NO: 64:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

GTCCTTAAGA AGGATGAAGG AGACCGCCG

29

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

TCGGGATCCT TAGCTGTCCA TGTGCTG

27

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

TCGGGATCCT CATTGTTTGC CTCCTG

27

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

TGCTCTAGAT CTTAACATGA AGAATGTTTT AG

32

[illegible]

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

TCGGATCCGC TTAAGTGGAA GCTTCCAAAC

TCGGATCCGC TTAAGTGGAA GCTTCCAAAC